

Pragmatic Requirements for Adaptive Systems: a Goal-Driven Modelling and Analysis Approach

Felipe Pontes Guimarães^{*}, Genaina Nunes Rodrigues[†], Raian Ali[‡] and Daniel Macêdo Batista^{*}

^{*}University of São Paulo, Brasil

Rua do Matão, 1010, São Paulo, SP – Brasil

{felipepg,batista}@ime.usp.br

[†]Universidade de Brasília, Brasil

Asa Norte, Brasília, DF CEP: 70910-900

genaina@cic.unb.br

[‡]Bournemouth University, UK

Fern Barrow BH12 5BB

rali@bournemouth.ac.uk

Abstract—Goal-models (GM) have been used in adaptive systems engineering for their ability to capture the different ways to fulfill the requirements. Contextual GM (CGM) extend these models with the notion of context and context-dependent applicability of goals. In this paper, we observe that the interpretation of a goal achievement is itself context-dependent. Thus, we introduce the notion of Pragmatic Goals which have a dynamic satisfaction criteria. We also developed and evaluated an algorithm to decide the Pragmatic CGM’s achievability. Finally, we performed several experiments to evaluate and to compare our algorithm against human judgment and concluded that the specification of context-dependent goals’ applicability and interpretations make it hard for domain stakeholders to decide whether the model covers all possibilities, both in terms of time and accuracy, thus showing the importance and contribution of our algorithm.

I. INTRODUCTION

Goal-Models (GM) are well established requirements engineering tools to depict and break-down systems using socio-technical concepts like inter-dependent actors, goals, quality goals or softgoals, tasks and resources [1]. GM facilitate the understanding of the system as a whole and provide a rationale about why the system needs to execute certain functionality and their possible variations [1]. In other words, it provides the goals for which the system should be designed and the various possible ways to reach those goals.

The variability of goal achievement strategies is the baseline for an actor to adapt by deciding which alternative to adopt as a response to certain triggers or adaptation drivers, e.g. faults, errors, availability of computational resources and newly available services and packages. The dynamic environment in which the system operates, *i.e.* its context, could also be an adaptation driver. The Contextual Goal Model (CGM) [2] extends the traditional goal model [3], [4] with the notion of context. Context may be an activator of goals, a precondition on the applicability of certain alternatives to reach a goal and a factor to consider when evaluating the quality provided by each of these alternatives.

However, we advocate another effect of context on CGMs and requirements in general. The interpretation of a goal

achievement is itself context dependent. This means that, in certain contexts, the mere achievement of the sub-goals in a goal model does not imply that the parent goal has been achieved. As an example, consider an ambulance dispatch. The goal of arriving at the patient’s location in timely fashion would be seen as achieved when this takes 15 minutes and he/she suffers from dizziness. However, the same goal would not be achieved if the patient suffered from a heart condition. The pragmatism, *i.e.* dynamic interpretation, is not about the quality but the boolean decision whether a goal is achieved.

In this paper, we introduce the concept of *pragmatic goals* to grasp and model the idea that a goal’s interpretation varies according to context. We define the achievability of pragmatic goals as being the capability of fulfilling a goal as interpreted within the context of operation. We also develop and implement an algorithm to compute the execution plan which is likely to achieve a pragmatic goal in a certain context.

We evaluate the applicability of our modeling and the necessity for a reasoning algorithm by applying it on a case study of a Mobile Personal Emergency Response System and comparing the performance and reliability of the answers generated by human volunteers and by our algorithm. Results showed that volunteers took up to 17 minutes to provide answers with 27.81% reliability whereas our algorithm provided correct answers in just a few milliseconds. Finally, we performed a scalability analysis to show the usability of our algorithm in pinpointing context sets in which the CGM as a whole may become unachievable, as well as the possibility of using it to support runtime adaptation by laying out an execution plan which is likely to achieve the necessary constraints.

The paper is organized as follows. Section II presents the CGM concept on which our model is based. Section III presents the pragmatic goals and pragmatic goal achievability concepts. Section IV presents the proposed model and automated reasoning to decide the pragmatic achievability. Section V demonstrates the applicability of the modeling and analysis approach. Section VI presents related work and Section VII concludes the paper and outlines our future work.

II. THE CONTEXTUAL GOAL-MODEL

Contextual Goal Model, proposed in [2], explicitly captures the relation between the goal model and dynamic environment of a system. It considers context as an adaptation driver when deciding the goals to activate and the alternatives - subgoal, task or delegation - to adopt and reach the activated goals. Context can also have an effect on the quality of those alternatives and this is captured through the notion of contextual contribution to softgoals. In terms of syntax and modeling constructs, context can be correlated to certain variation points in the goal model. It is also analyzed through a technique called Context Analysis which allows to derive a formula, made of observable pieces of information (facts). This formula represents, in a measurable way, the condition whether a context holds.

Context is defined as the reification of the system's environment, *i.e.*, the surrounding in which it operates [5]. For goal models, context is defined as a partial state of the world relevant to an actor's goals [2]. An actor is an entity that has goals and can decide autonomously how to achieve them. A context may be the time of a day, a weather condition, patient's chronic cardiac problem, etc.

The CGM presented in Figure 1 depicts the goals to be achieved by a Mobile Personal Emergency Response System which is meant to respond to emergencies in an assisted living environment. The root goal is "respond to emergency", which is performed by the actor *Mobile Personal Emergency Response*. The root goal is divided into 4 subgoals: "emergency is detected", "[p] is notified about emergency", "central receives [p] info" and "medical care reaches [p]" ([p] stands for "patient"). Such goals are then further decomposed, within the boundary of an actor, to finally reach executable tasks or delegations to other actors. A task is a process performed by the actor and a delegation is the act of passing a goal on to another actor that can perform it.

For instance, the goal "setup automated [p] info" is decomposed into two subgoals: "[p] location is identified" and "[p] situation data is recovered". Such subgoal may then be realized via the task "access data from database". An example of delegation can be seen at the goal "ambulance is dispatched to [p] location". This goal is not performed by the *Mobile Personal Emergency Response* system but rather delegated to another actor which is the *Ambulance Dispatching System*.

The CGM observes that not all the subgoals, delegations and/or tasks are always applicable. Some of them depend on certain contexts whether they hold. For instance, the task "identifies [p] location [l] by voice call" is only applicable if there is GSM (Global System for Mobile Communications) coverage at [p]'s location. The part of the CGM in which only applicable refinements, or nodes, remain is called the CGM variant for the current context. The possible contexts (C1-C12) are listed in the lower side of Figure 1.

III. PRAGMATIC REQUIREMENTS

Traditionally in a CGM, achieving one (OR-Decomposition) or all (AND-Decomposition) of the subgoals is seen as a satisfactory precondition for achieving the parent goal. We argue that the achievement of some goals would need to be seen from a more pragmatic point-of-view and not as a straightforward implication of the achievement of other goals or the execution of certain tasks. The decision whether a goal is achieved could be context-dependent. Thus we need a more flexible definition of goals to accommodate their contextual interpretation and achievement measures.

The representation of the quality of achievement of a goal as a softgoal is different from the pragmatism of the goal achievement. The pragmatic nature of a goal is not a matter of achieving it with higher or lower quality, but achieving it at all. Also, it has to do with the context at the time of execution and not with the model itself, making the quality requirements more strict or even relaxing them when some contexts apply.

Take the example of Figure 1: in general, the ambulance may take up to ten minutes to arrive. However, for a patient with a minor discomfort it can take its time and arrive about ten minutes later without suffering any penalty. For this situation, there is no need to hurry since there is no life threat. On the other hand, if the ambulance takes the same ten minutes to reach a patient having a heart attack, one cannot say the goal was achieved. In these situations, the delivered level of quality may not be a separate part from the boolean answer of whether a goal is achieved or not. Such quality level is intricately integrated, in a quantifiable way, into the very definition of the goal's achievement.

To be able to illustrate the pragmatic nature of some goals, let us consider the CGM from Figure 1 and focus on the highlighted goal "[p] location is identified" (this goal will be called G_{loc}). The CGM states that there are up to four ways of achieving this goal: either by considering the last known location of the patient, by a voice call asking the patient on his/her location, by GSM signal triangulation, or by GPS lock. In the traditional goal model interpretation, given that at runtime the proper contexts hold, the execution of any of these tasks is enough to consider G_{loc} achieved. In this conception, G_{loc} is a clear example of hard goal.

However, from a pragmatic point-of-view, the simple fact that the location was discovered may not be enough. A location estimate based on the GSM signal triangulation may have a very low precision, a voice call may be achieved but the patient may not know how to describe its position, and a GPS lock may take too long in some cases. These nuances could be accepted in certain contexts while in other contexts they may lead to consider the goal unsatisfied. As a result, the pragmatic requirements must come into play.

A pragmatic goal describes the means to achieve it and also the interpretation of that achievement. This interpretation, which depicts the goal's pragmatic fulfillment criteria, can be expressed as a set of mandatory and crisp, therefore quantifiable, quality constraints (QCs). Unlike softgoals, these

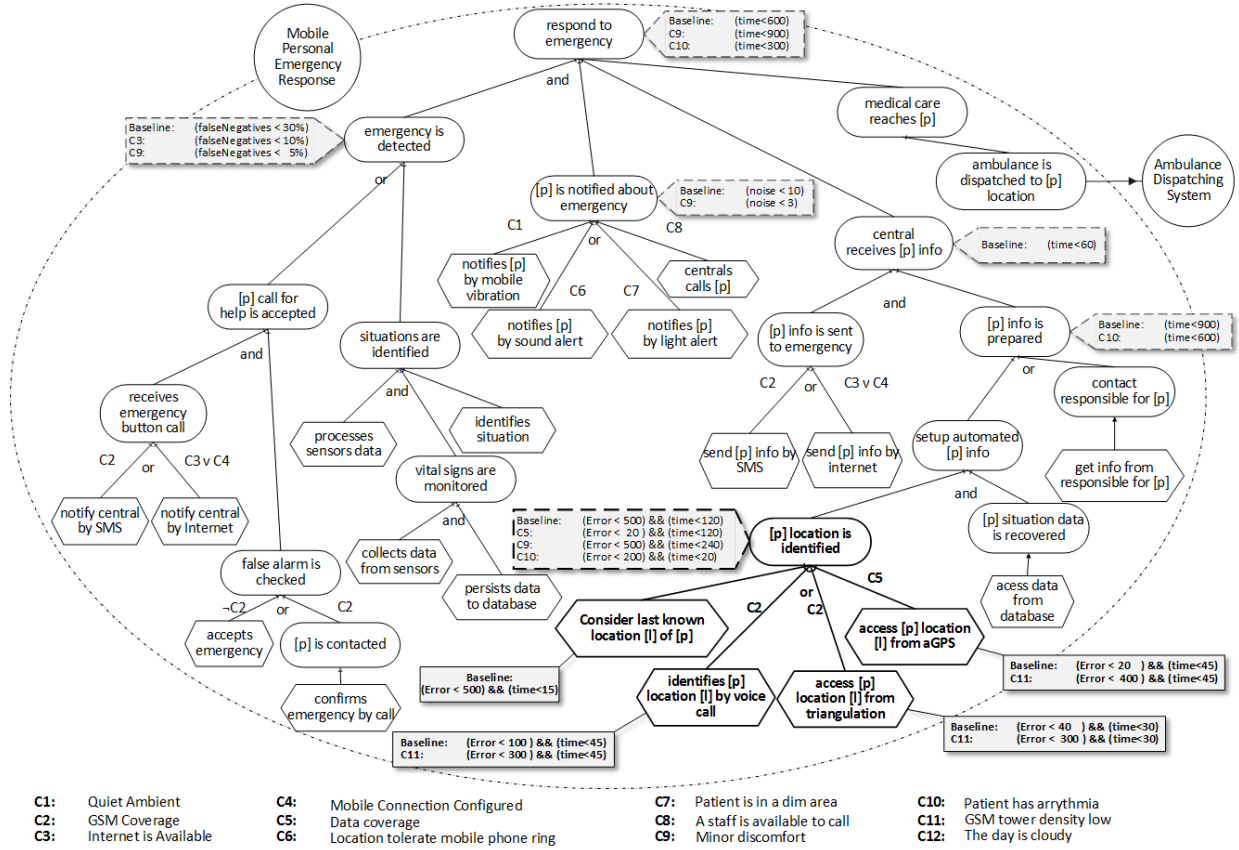


Figure 1. A CGM for responding to emergencies in an assisted living environment (adapted from [6])

QCs are quantified measures needed for the fulfillment of a goal and an inherent part of its definition. In the previous example, the goal G_{loc} could then be defined as: “in order to reach goal G_{loc} , the location must be identified so within an error radius of maximum 500m and in less than 2 minutes”. Again, this would not suffice, as a radius of 500m and 2 minutes might be an over-relaxed condition for patients under critical conditions. On the other hand, setting the highest possible level of demand for all situations is likely unreal and could lead to a huge waste of resources.

This brings into light another aspect to be taken into account for the pragmatic requirements: the fact that the interpretation for the achievement of a goal is itself context-dependent. We consider that there is a basic default condition for achieving a goal. On top of that, and for specific contexts, we could relax or further strengthen condition which interprets whether a goal is achieved. We propose that the contextual QCs on the achievement of a goal should be captured together with the other effects of context in the CGM. One advantage of capturing the pragmatic goals within the CGM is to enable reasoning on the possibility of achieving a goal under the current context and current constraints. We differentiate these interpretations in the sense that a relaxation condition is not mandatory but a condition that further strengthen the QC must necessarily be considered.

For instance, in the previous example, a QC of getting a location within 500m in less than 2 minutes is a default constraint. However, if the user has access to mobile data connection (context C5) then a much preciser location can be obtained from the GPS. Under these circumstances, a lock within 500m may seem like an over-relaxed constraint. For a patient with cardiac arrhythmia (context C10), a more strict QC is needed. Suppose that the system has to ensure that an ambulance reaches the patient’s home within 5 minutes. Possibly, in this case, a faster but less precise location would be better suited. Such nuances in the interpretation of the goal are summarized in Table I. In the three specific contexts, the interpretation must be different than the baseline. For example, the requirements for a minor discomfort (context C9) are more flexible than the requirements for an arrhythmia (C10). In this case, the interpretation of goal G_{loc} may be expressed by Table I and also presented as a box near the goal “[p] location is identified” on Figure 1.

To be able to differentiate traditional goals from goals where the delivered quality of service (QoS) defines the condition for the goal’s achievement, we introduced pragmatic goals into the CGM. A pragmatic goal is a hard goal with an interpretation expressed as a set of context-dependent QCs, shown in the graph within the dashed boxes. Unlike quality attributes and softgoals, a QC describes a *mandatory* condition

Table I
QUALITY REQUIREMENTS FOR SUCCESSFULLY ACHIEVING
[P] LOCATION IS IDENTIFIED

Context	Interpretation	
Baseline	(error < 500m)	&& (time < 120s)
C5	(error < 20m)	&& (time < 120s)
C9	(error < 500m)	&& (time < 240s)
C10	(error < 200m)	&& (time < 20s)

for considering a goal achieved. Such QCs may be imposed by the goal itself, i.e. by definition, as well as by each individual actor to reflect its own interpretation of the goal's fulfillment in the different contexts.

A. Achievability of Pragmatic Goals

Not only a goal's fulfillment interpretation may vary according to the context, the QoS levels delivered by the tasks, i.e. executable processes, may also themselves differ. The same task may provide different QoS when executed in different contexts. This is represented by the boxes linked to tasks¹. This will propagate and affect the overall provided QoS of the parent goals.

The reasoning part of our algorithm (Subsection IV-A) considers that pragmatic goals can only be achieved if their provided QoS comply with the QCs specified for them where both are context-dependent. This means that we extend the basic effect of context on a CGM to cover success and achievement criteria. Such expressiveness enables further analysis for a key adaptation decision: how to reach our goals while respecting the QCs under the current context where the goals interpretation, the space of applicable alternatives to reach a goal and the QoS provided by the tasks are all context-dependent.

This part also considers the situation where it may not be possible to meet the QoS standards which meet the goal's interpretation through any of the applicable sub goals, tasks and/or delegations as they deliver not a static but a context-dependent QoS level. In such cases, we classify the goal as unachievable and the reasoning part can explain the reason.

To explain our rationale, let us consider the goal G_{loc} and the contexts impacting on its interpretation. In the presented CGM, given any moment in which $\neg C5 \wedge \neg C9 \wedge \neg C10$ holds, the goal [p] location is identified is interpreted as met as long as the error margin is less than 500m precision and the location is given within 120 seconds. However, when $C5 \wedge C9 \wedge C10$ holds, such quality is insufficient and the required quality standards to meet this goal are more strict. In this case, the error margin cannot be more than 20m and must be given within 20 seconds. Finally, when $\neg C2 \wedge \neg C5 \wedge C10$ holds, it would not be possible to achieve this goal because the only applicable refinement left would be considering last known location of [p]. This option, however, has a much higher error margin than the required 200 meters.

¹Due to space limitations only the four tasks under G_{loc} have their context dependent apparent. This does not mean they do not exist in the other tasks

In the above example, the conclusion is that under a certain context the system may not be able to determine the patient's location with the required precision. This, in practice, does not mean doing nothing. The motivation to do this analysis is because having such knowledge beforehand would allow consideration of other strategies, like adding more alternatives to the same goal to cover a larger range of contexts. At runtime, this conclusion would lead to search for a better variant at a higher-level goal by choosing another branch of an OR-decomposition, which is able to deliver the required quality standard. Therefore, our analysis is both meant for design-time - reasoning to evaluate and validate the comprehensiveness of the solution - and for runtime - searching for the right alternative to reach goals in a specific context.

IV. PRAGMATIC GOAL MODEL

In this section, we concretize our extension to the CGM and elaborate on the new constructs we add as well as their semantics. We mainly enhance the CGM with context-dependent goal interpretations and the expected delivered QoS, which are also context-dependent, for tasks in order to reason about the achievability of the goals for which these tasks are executed.

To model pragmatic goals and their interpretation, we have extended the CGM's concept of a goal. A goal is refined into subgoals, tasks and/or delegations which must be achieved or performed to meet the parent goal. We extended this concept so that a pragmatic goal can now have an interpretation in the form of QCs that have to be met in order to render the pragmatic goal achieved at any given context. Thus we provided a quantifiable measure for a goal which encompasses the verifiable satisfaction criteria and their dynamics in the different contexts.

Figure 2 presents a conceptual model of our extension to the CGM. For the focus of this paper, the CGM could be seen as an aggregation of Refinements. A Refinement may be specialized into several types: Tasks, Delegations and Goals. A Delegation represents when the Goal is pursued not by the current but by an external actor. Tasks are performed by the actor in order to achieve a goal. Tasks may report the expected delivered quality for each metric through the providedQuality method. Goals have a Refinements set which define the subgoals, tasks and/or delegations that can be used for achieving it as well as a method to distinguish AND- from OR-compositions.

Pragmatic Goals extend the Goal concept with an Interpretation. A goal Interpretation is an abstract concept that has the function of cross-referencing a context and the appropriate Quality Requirement for that given context. The pragmatic goal is said to be achieved *iff* such requirements are met. Otherwise, the goal's delivered QoS is considered inappropriate and the goal is not achieved regardless of achieving one or all of its Refinements.

The Quality Constraints are expressed in terms of the applicableContext in which it holds, the metric

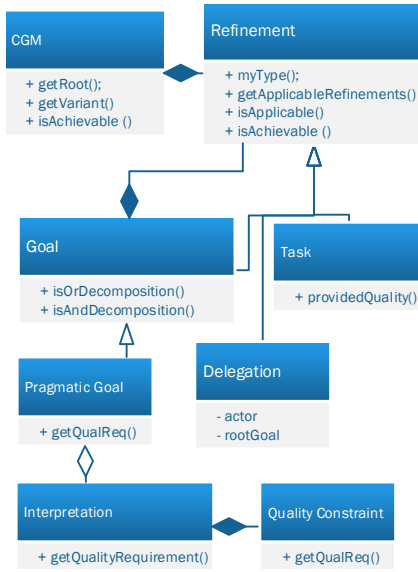


Figure 2. Conceptual metamodel for our extension to CGMs

that should be considered, the `threshold` which is a numerical value that represents the scalar value for such metric and the comparison which defines whether the threshold described is the maximum allowed value or the minimum. For instance, to state a quality requirement of at most 250ms for the execution time when context $C1$ holds, the `metric` would be “ms”, `threshold` would be 250, `condition` would be “Less or Equal” and `applicableContext` would be $C1$.

Every `Refinement` inherits the `isAchievable` method. This method can be used either by the final users or by the higher level goals to define whether a particular goal can be achieved for a given quality requirement under the current context. Intermediate goals also have their own predefined `Interpretation`. While this is obviously necessary for the root goal, as the ultimate objective, we also allow certain subgoals to be defined as pragmatic. In principle, actors should be able to impose further constraints on the criteria for achieving any goal within their boundary. The importance of the subgoals quality requirement becomes obvious when dealing with delegation of goals where the external actor may have itself a different, more relaxed or more strict, quality constraint, not necessarily compatible with what the delegator intends.

Both the expectation of delivered quality by the tasks and the quality constraints for the goals, subgoals or delegations are added to the CGM. This is meant to be done by the requirements expert or the domain experts due to the need for specialized knowledge to define such metrics. Different organizations may have different definitions of these interpretations and constraints which is yet another facet of the pragmatism of our approach.

A. Achievability evaluation method

In this section we revisit the classical concept of achievability of a goal to fit the nature of Pragmatic CGMs. On top of the basic context effect on a CGM, we enable a

higher model expressiveness. Such expressiveness will enable richer adaptation decisions which not only consider the static achievability but also the achievability under the dynamic context and its effect on the fulfillment criteria of a goal. The achievability of a goal and the space of adoptable alternatives to achieve it are essential information to plan adaptation, seen as a selection and enactment of a suitable alternative to reach a goal under a certain context.

To evaluate the achievability of a particular pragmatic goal we present the algorithm in Figure 3. It implements the `Refinement` entity’s `isAchievable` method (Figure 2) and correlates three context-dependent aspects from the model: (1) the applicable refinements; (2) the goals’ interpretations and; (3) the delivered quality level provided by the tasks.

Require: CGM, current context and desired QCs

```

1: Goal root ← cgm.getRoot()
2: if !root.isApplicable(current) then
3:   return NULL
4: end if
5: if (root.myType() == task) then
6:   if (root.canFulfill(qualReq)) then
7:     return new Plan(root)
8:   else
9:     return NULL
10:  end if
11: end if
12: QualityConstraint consideredQualReq
13: consideredQualReq ←
   root.interp.stricterQualityConstraint(root.qualReq,
   qualReq)
14: Plan complete ← NULL
15: deps ← root.getRefinements(cgm, curContext)
16: for all Refinement d in deps do
17:   Plan p ← d.isAchievable(cgm, context,
   consideredQualReq)
18:   if (p != NULL) then
19:     if (root.isOrDecomposition()) then
20:       return p
21:     end if
22:     if (root.isAndDecomposition()) then
23:       complete ← addPlanToPlan(p, complete)
24:     end if
25:   else if (root.isAndDecomposition()) then
26:     return NULL
27:   end if
28: end for
29: return complete

```

Figure 3. `isAchievable(CGM cgm, Context current, QualityConstraint qualReq)`

The algorithm decides whether the root goal is achievable and, if so, lays out an execution plan, i.e., the set of all tasks to be executed, likely to achieve the desired QCs. The algorithm is recursive with a proven linear complexity with respect to

the number of refinements in the CGM², building on the fact that the CGM is a tree-structured model without loops and that each refinement may be seen as a tree node.

The algorithm considers the root node of the CGM (line 1) and checks whether the root goal is itself applicable under the current context (line 2), returning NULL if it is not (line 3). In the particular case when the variant's root node is a task (line 5) it can readily decide on the achievability. This is because the task nodes know the expected QoS it can deliver for each metric under the context considered in the CGM. By comparing the delivered QoS and required QCs (line 6), the node can decide whether it is capable or not of delivering such QCs. If it can, it will return a plan consisting only of this task (line 7), otherwise it will return NULL (line 9) and indicate its inability to fulfill the goal's interpretation.

If the root is not a Task, the algorithm will define its quality requirement as the stricter *Quality Constraints* between its own and the QCs passed on as parameters (line 13) and begin laying out an execution plan to fulfill such QCs (line 14). For each of the applicable refinements, it will evaluate if it is achievable (line 17). If the refinement is achievable then, for OR-decompositions, the algorithm returns this plan immediately (lines 19 - 20) and for AND-decompositions it is added to the *complete plan* (lines 22-23). Otherwise, if the refinement is unachievable it will immediately return NULL for AND-decompositions (line 26). Finally, for AND-decompositions, should all refinements are achievable it will return the *complete plan* (line 29).

As an outcome, an execution plan is returned for achievable goals. For unachievable goals the NULL value is returned to indicate the inability of fulfilling the required constraints, allowing for alternate means of achieving higher level goals to be explored.

V. PRAGMATIC MODEL AND ACHIEVABILITY ALGORITHM EVALUATION

In this Section, we aim to show the need for an algorithmic approach to handle Pragmatic Goals and to evaluate the capability of the proposed model to scale over the Pragmatic CGM size with regard to the amount of goals and contexts.

To do so we used the Goal-Question-Metric (GQM) evaluation methodology [7]. GQM is a goal-oriented approach used throughout software engineering to evaluate products and software processes. It assumes that any data gathering must be based on an explicitly documented logical foundation which may be either a goal or an objective.

GQM's first step is to define high-level goals to be evaluated. For each goal, a plan consisting of a series of quantifiable questions is devised to specify the necessary measures for duly assessing the evaluation [7]. These questions identify the necessary information to achieve the goals while the metrics define the operational data to be collected to answer each question.

²Formal demonstration available at <https://github.com/felps/PragmaticGoals/blob/master/AlgorithmComplexity.pdf>. Accessed on January 22th, 2015

In such a methodology, the main goals of our evaluation are: (I) to show the need for an algorithmic approach; (II) to evaluate the capability of using our approach for adaptive systems at runtime; (III) to evaluate whether our approach across may be used to identify context combinations which render the Pragmatic CGM's root goal unachievable by construction.

From these goals, the GQM plan was defined and is presented in Table II.

A. Experiment Setup

The experiment setup consisted in two parts: the human capability for evaluating a Pragmatic model and the algorithm scalability analysis. These parts and their evaluations were engineered to provide the metrics demanded by the GQM plan (Table II)

For the human capability evaluation, we performed an experiment with 55 volunteers from the fields of computer science, software, electronics and automotive engineering. The group would be given a 10 minutes explanation on the concepts concerning the Pragmatic CGM and a 5 minutes presentation of the CGM from Figure 1, its 6 pragmatic goals (up to three different interpretations) and the provided QoS for all tasks (up to 3 context-dependent values).

After the explanation, the volunteers received 4 context sets (sets 1,3 and 4 were achievable while set 2 was not) and were asked to identify a set of tasks that could fulfill the CGM's pragmatic aspect under that context set whenever the goal is achievable or, otherwise, state the goal as unachievable. They were instructed to check the wall clock and write down the time once they came up with the solution.

To limit the amount of necessary volunteer time, a limit of 25 minutes was set. After such deadline there would be no need to continue since we considered that taking more than 6 minutes to decide on a single context set for a rather small model would already be inadmissible and make it impossible for human judgment to be considered as a plausible alternative for finding unachievable scenarios.

As for the algorithm evaluation, all experiments to evaluate the necessity, correctness and performance of the algorithm were implemented as automated tests under Java's JUnit framework³. This guarantees that the evaluation is both effortless and repeatable.

Each time measurement was performed 100 times over each model and the amount of time was measured using Java's `System.nanoTime()` feature. This method provides an approximate measure of time with precision of at least 1 ms. The final measurement was the average of all 100 executions.

The context set coverage evaluation was performed similarly, but instead of executing 100 executions, it would continuously execute until all the context sets were covered or until the elapsed time surpassed ten seconds.

The algorithm from Figure 3 was implemented⁴ using Java

³The evaluation mechanisms, the complete result set as well as the implementing code are available at <https://github.com/felps/PragmaticGoals>. Accessed on January 22th, 2015

⁴Source code is available at <https://github.com/felps/PragmaticGoals>. Accessed on January 22th, 2015

Table II
GQM DEvised PLAN

Goal 1: Determine the necessity for an algorithmic approach	
Question	Metric
1.1 How long would a human take to come up with an answer?	Time to produce an answer
1.2 How reliable would an answer provided by a human be?	Percentage of correct answers
1.3 How long would the algorithm take to come up with an answer?	Time to produce an answer
1.4 How reliable would an answer provided by the algorithm be?	Percentage of correct answers

Goal 2: Evaluate the algorithm's runtime usage capability	
Question	Metric
2.1 How does the algorithm scale over the amount of goals in the model in average?	Execution time
2.2 How does the algorithm scale over the amount of contexts in the model in average?	Execution time
2.3 How does the algorithm scale over the amount of goals in the model in the worst case scenario?	Execution time
2.4 How does the algorithm scale over the amount of contexts in the model in the worst case scenario?	Execution time

Goal 3: Evaluate the algorithm's capability of pinpointing unachievable context sets	
Question	Metric
3.1 Can the algorithm cover all context sets for models with increasingly large models and reasonable context amounts?	Context sets coverage

OpenJDK 1.7.0_65 and the evaluation tests were performed on a Dell Inspiron 15r SE notebook equipped with a Intel Core i7 processor, 8GB RAM running Ubuntu 14.10, 64 bits and kernel 3.16.0-29-generic. We also used the EcLeMma Eclipse's plugin for ensuring the tests' code coverage.

B. Goal 1: Determine the necessity for an algorithmic approach

We advocate that the CGM, though very useful in sharing the view and understanding of the problem among system developers and stakeholders, is just too complex to be merely evaluated by human judgment. On the other hand, it is very fast and very simple for a computer to perform a similar analysis and present reliable and comprehensive data.

We conducted an experiment with volunteers from the University of Brasilia and compared to the algorithmic approach performance.

1) **Questions 1.1 and 1.2 - The human perspective evaluation:** In order to answer questions 1.1 and 1.2, we performed an experiment with 55 volunteers in the fields of computer science, software, electronics and automotive engineering. The participants were familiar with the use of modeling, even if not necessarily within a software engineering environment but also in the automotive and product design. This added weight to our experiment as the practitioners of our approach are also those who are product designers.

The group had an explanation on the concepts concerning the Pragmatic CGM and a presentation on the CGM from Figure 1, its 6 pragmatic goals and the provided QoS for all tasks (up to 3 context-dependent values). After that, any remaining doubts about the Pragmatic CGM concept and/or the CGM itself were clarified.

After the explanation, the volunteers received 4 context sets (sets 1,3 and 4 were achievable while set 2 was not) and were asked to identify a set of tasks that could fulfill the CGM's

pragmatic aspect under that context set whenever the goal is achievable or, otherwise, state the goal as unachievable.

For this experiment, we measured the time the volunteers took to produce an answer for each set (question 1.1) and the correctness of the answers produced (question 1.2).

a) **Question 1.1: How long does a human take to come up with an answer?:** The average time for a volunteer to produce a solution for each of the given context sets of the experiment is shown in Figure 4. The box plots represent the median and the dispersion of the time (in seconds) it took the volunteers to come up with an answer for each provided set. These times varied a lot since it took a while for the participants to understand the whole idea and, as they progressed through the sets they gained experience thus becoming faster. Still in the best case (Context Set 4) most of them took between 100 and 180 seconds to come up with an answer.

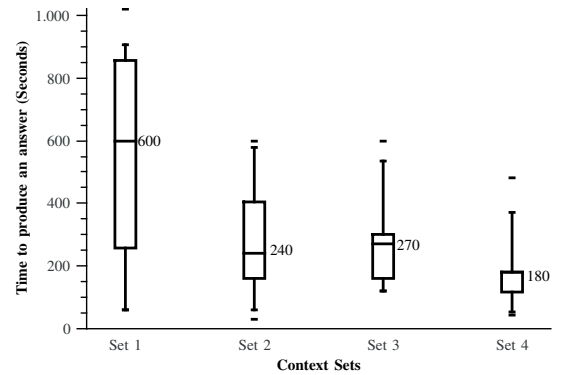


Figure 4. Average Time (in seconds) for volunteers to produce an answer to each context set

b) **Question 1.2: How reliable would an answer provided by a human be?:** All the collected answers from the participants were then compared to the correct options. Out of the

97 answers produced, 54 were considered erroneous, 13 were partially correct and only 26 were precise. This means that 73.19% of the provided answers were only partially correct or completely incorrect.

Our experiment also showed that it was easier for the participants to correctly state a CGM as unachievable (75% correct answers for set 3) than to find a task set which would satisfy the Pragmatic Requirements (11.6% correct answers for sets 1, 2 and 4). This is probably because deeming a subgoal as unachievable may propagate this condition to the whole CGM or at least to an AND-decomposition sub-tree whereas finding a valid solution for the whole tree requires thoroughly investigating all options.

2) Questions 1.3 and 1.4: The algorithmic approach evaluation: The implemented algorithm receives three arguments as input: the CGM, the set of contexts and an optional user's quality constraint for the CGM's root goal. It outputs NULL if the root goal is unachievable or an execution plan, *i.e.*, a set of tasks that can achieve the root goal. The execution plan must abide both by the quality constraint provided as input as well as any pragmatic goals' interpretation.

a) Question 1.3: How long would the algorithm take to come up with an answer?: To evaluate the time for the algorithm execution on the CGM of Figure 1, we executed 1000 iterations of the algorithm for each context set. The results showed that the algorithm took, in average, less than 1 ms to be executed in each of the 4 scenarios.

b) Question 1.4: How reliable would an answer provided by the algorithm be?: To validate the algorithm's correctness, we implemented tests for each context set used with the volunteers. For each one of these, we have identified all of the inapplicable tasks - both because of context or quality constraints - and asserted that the outputted execution plan did not contain any of these. All the test succeeded thus providing evidence of the algorithm correctness.










Element	Missed Instructions	Cov.	Missed Branches	Cov.
Refinement		100%		100%
Interpretation		100%		100%
Goal		100%		100%
Plan		100%		n/a
Pragmatic		100%		n/a
CGM		100%		n/a

Figure 5. Eclipse's EclEmma plugin reporting 100% code coverage for main classes⁶

To further evaluate the correctness of the algorithm, we have also implemented over 70 test cases for the whole implementation. These were sufficient to achieve 95.2% overall code coverage. In particular, we paid special attention to the Goal, Pragmatic Goal and Refinement classes as well as to the `isAchievable` method which were extensively tested until achieving 100% code coverage and is presented in Figure 5.

C. Goal 2: Evaluate the algorithm's runtime usage capability

One of the purposes of this algorithm is to enable the layout, at runtime, of an execution plan which is able to achieve the CGM's root goal. To do so, the algorithm needs to be able to process models with varying complexities, both in terms of context amount and CGM model size, in a reasonable amount of time so that it won't seriously impact the response time.

c) Question 2.1 and 2.2: How does the algorithm scale over the amount of goals and contexts in the model in average?: To evaluate the algorithm's scalability over the model size in terms of goals and contexts amount, we implemented the following test: for each combination of CGM model size (100 to 10000 nodes in steps of 100 nodes) and amount of contexts (1 to 20 contexts), the test would randomly generate 100 CGM models and then the `isAchievable` method was executed 100 times in each model and the average execution time was measured. Finally, it outputs the average execution time for each combination. The resulting average times are presented in Figure 6a.

d) Question 2.3 and 2.4: How does the algorithm scale over the amount of contexts in the model in average and in the worst case?: Similarly to the previous experiment, to evaluate the algorithm's scalability in the worst case we implemented another test. This time the generated model would be the algorithm's worst case scenario: an achievable Pragmatic CGM composed solely of AND-Decompositions. This forces the algorithm to traverse the whole tree. The test generated random Pragmatic CGMs with sizes varying from 100 to 10000 nodes and, for each model, performed 100 executions. The observed average time per execution is shown in Figure 6b. The results were similar in behavior though higher than those observed in the average case.

D. Goal 3: Algorithm's capability of pinpointing unachievable context sets

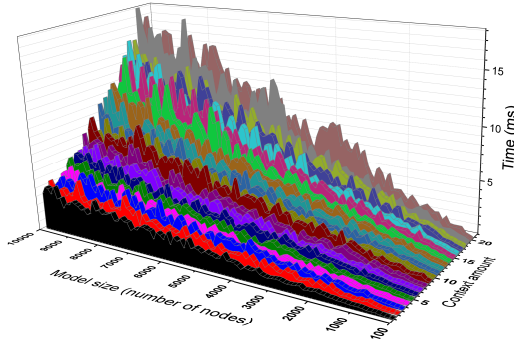
To answer question 3.1 (How many contexts sets can the algorithm evaluate per second for increasingly large models?) we implemented one last test which would generate increasingly larger models with a fixed set of 15 contexts. For each model size 10 random models were generated. Finally, on each of these models and for each combination of the 15 contexts, we have executed the algorithm and measured the percentage of possible context sets it was able to sweep, either finding a suitable solution or not, within ten seconds.

The results can be seen in Figure 7. As it shows, on smaller models up to 300 goals, the algorithm was able to fully sweep the 32768 context sets within the stipulated deadline. On larger models - up to 5000 goals - the algorithm was able to sweep around 40% of the combinations. Even for rather large models with 10000 nodes it was able to cover more than 25% of the possible combinations.

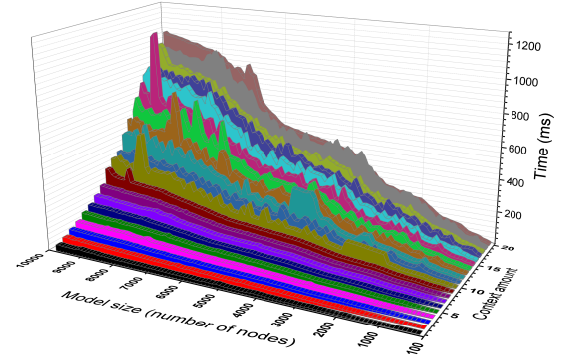
E. Discussion of the results

As stated in the GQM plan (Table II), the experiments had three main goals: (1) determine if there is the necessity for an algorithmic approach, (2) determine if the algorithm could be

⁶<http://www.eclEmma.org/>. Accessed on January 22th, 2015



(a) Average case



(b) Worst case scenario

Figure 6. Algorithm's scalability over the model size, in number of nodes, and context amount

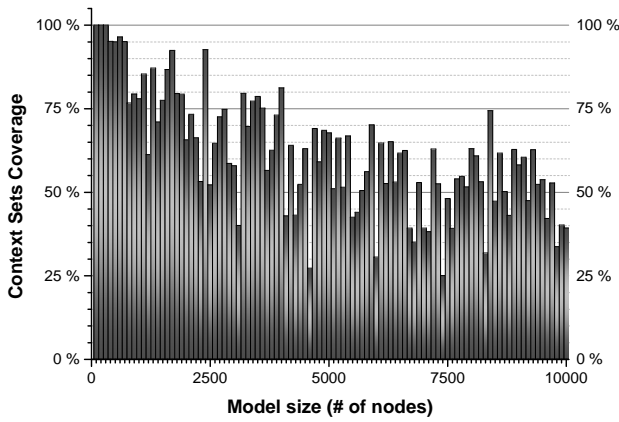


Figure 7. Percentage of possible context sets swept within 10 seconds versus model size

used at runtime to find a suitable plan for the current context and; (3) whether it could also be used for pinpointing context combinations in which there is no applicable goals/tasks combination sufficiently good to satisfy the pragmatic goals' requirements.

For the first goal, the results have corroborated with the understanding that the Pragmatic CGM complexity is far too great to be dealt simply by the human perspective. On the other hand, the algorithm proved itself as a much faster and deterministic alternative. While the volunteers took up to 17 minutes to provide an answer with 73.19% reliability, the algorithm produced an answer to the same problem in under 1 millisecond. Also, since it is a deterministic algorithm, the produced answers are always valid. Even if considering very large improvements on the human performance, the algorithmic approach would, most likely, still largely surpass human performance.

Though the comparison between volunteers and an algorithm may seem unfair, the goal of this comparison is to determine the necessity of an algorithmic approach. Being so and as expected, the algorithm has proved itself much better both in terms of efficiency and efficacy as well as subsiding the desired hypothesis that an algorithm is indeed necessary.

For the second goal, we aimed to evaluate whether the algorithm was suitable for execution at runtime. This meant verifying two main aspects: that the algorithm would efficiently execute even with large Pragmatic CGM models and within a reasonable time to not affect the response time. We performed this analysis for random models (Figures 6a and 4) which may have incurred in the observed high variability for the average case. However, an upper boundary was also set with the worst case evaluation (Figure 6b). With regard to the first aspect, Figure 6 shows that the algorithm's execution time grows linearly over the amount of goals as well as over the amount of contexts in the Pragmatic CGM model, both in the average and in the worst case scenarios. The second aspect can also be derived from Figure 6. As it shows, even when considering the worst case scenario, the time for evaluating a model with 10000 nodes and 20 contexts was 1081 ms. As a matter of comparison, the default setting of Axis2 (Java's API for web services) for a read timeout is set to be 300 seconds⁷. Therefore, we see this as an acceptable result for the purposes of showing that the algorithm's performance is enough not to severely impact runtime.

Finally, for the third goal we evaluated how long the algorithm took to sweep all context combinations. Given the results from Figure 7, we can state that for models with up to 10000 goals and 15 contexts, it is actually possible to evaluate all context set combinations within one minute. Thus, the proposed algorithm can indeed be used on a Pragmatic CGM to pinpoint unachievable scenarios with up to 15 contexts, which can then be scrutinized by the CGM designer in order to correct or document it.

F. Threats to validity

Construct validity concerns establishing correct operational measures for the concepts being studied. This was minimized by the usage of the GQM methodology to lay out the evaluation plan. Firstly, the goals that needed to be achieved were laid out; for each goal we envisioned the questions which needed

⁷http://www-01.ibm.com/support/knowledgecenter/SSD28V_8.5.5/com.ibm.websphere.nd.doc/ae/rwbs_jaxwstimeouts.html. Accessed on January 14th, 2015

to be answered and only then the metrics were defined with these questions in mind.

Internal validity concerns establishing a causal relationship, whereby certain conditions are shown to lead to other conditions. The experiment setups guaranteed that all computer experiments and evaluations were performed in the same resource and environment. At each evaluation, up to two controlled variables were used.

External validity concerns establishing the domain to which the findings can be generalized. We conducted the assessment in the context of the Mobile Personal Emergency Response, which is specific to a given domain. Nevertheless, future work should assess this in different and real-life domains and with higher number of goals, contexts and quality constraints. In particular, this would help to further assess the scalability we observed. Furthermore, other non-functional properties could be evaluated and we consider it as future work.

VI. RELATED WORK

Previous work have tackled similar problems but to the best of our knowledge none has dealt with the dynamic context-dependent interpretation of requirements and, in particular, of goals. Relevant approaches include the work of Souza and Mylopoulos on Awareness Requirements which are goals that define quality objectives for other goals [8], [9]; the RELAX framework which provides a more rigorous treatment of requirements explicitly related to self-adaptivity, but it does so in a static, yet fuzzy, interpretation [10]; Baresi and Pasquale on Live Goals: goals whose individual behavior changed in order to pursue some qualitative objective and bring the system back to a consistent state [11], [12]; Dalpiaz et al. on declarative goals, which are separate goals whose achievement depends on the effects of its refinements on the environment [13] and; Sebastiani et al. deal with the Goal-Model satisfiability problem and its mapping into a propositional satisfiability (SAT) problem [14]. We argue that the notion of pragmatic goals could enrich the rationale of adaptation proposed in these and other approaches in goal-driven adaptation and that it differs from previous approaches since the above work considered it as a system- or model-wise problem instead of thinking it on a case-to-case context-dependent situation.

In comparison to the presented work, we differ from Souza and Mylopoulos who consider the quality objective as a distinct goal while we believe it to be an inseparable component of the goal itself: the mere completion of one or all refinements is not enough to achieve a goal, there may be clients' expectations/demands which must be met and which may vary over different contexts. In comparison to the RELAX framework, we differ in the sense that in our approach a goal's interpretation is not static but context-dependent. We also differ from Baresi and Pasquale in the sense that live goals would react to an inconsistent system state and reason over these particular goals' alternatives while our approach reasons over the whole CGM tree in an effort to, *a priori*, identify and avoid alternatives which will put the system in an inconsistent state, thus maximizing the probability of success. Finally, we

differ from Dalpiaz since we believe that the pragmatic nature is not a separate goal but intrinsically related to the goal itself and from Sebastiani since, to enable the algorithm's recursion and obtain a linear complexity algorithm, we use the simplifying assumption that there are no contributions or denials between different goals, thus enabling the treatment of the CGM as a tree and not as a generic graph.

The novelty of our work in comparison to other approaches in requirements-driven adaptation is twofold: (1) The definition of pragmatic goals which means that the satisfaction criteria for goals is context-dependent. (2) The development and implementation of an automated reasoning that can deterministically answer whether the goal is pragmatically achievable and, if it is, point out an execution plan that is likely to achieve it under the current context.

VII. CONCLUSIONS AND FUTURE WORK

In this paper we proposed the utilization of a Pragmatic CGM in which the goals' context-dependent interpretation is an integral part of the model. We have also shown why hard goals and soft goals are not enough to grasp some of the real-world peculiarities and context-dependent goal interpretations.

We defined the pragmatic goals' achievability property. A goal's achievability states whether there is any possible execution plan that fulfills the goal's interpretation under a given context. We also proposed, implemented and evaluated an algorithm to decide on the achievability of a goal and lays out an execution plan.

We compared the performance of our algorithm to that of a layman's analysis and effectively shown that an algorithmic approach to support the pragmatic goals is needed, considering that human judgment will probably not be fast nor reliable enough. Then, we discussed how the algorithm may enhance requirements engineering by evaluating and pinpointing context sets under which the root goal may not be achieved.

Finally, we performed a scalability analysis on it and shown that it scales linearly over the amount of goals and context amount. We have also shown that, for models up to 10000 nodes and 20 contexts, our algorithm is able to lay out an execution plan in about a second. Finally, we also evaluated that it was able to sweep all context combinations for models with 15 contexts and up to 10000 nodes in less than a minute, thus making this algorithm also suitable for pinpointing unachievable contexts.

For future work, we plan to: (1) integrate this algorithm into a CGM modelling tool; (2) study the possibility of the algorithm to return all task sets instead of a single one and (3) how to enhance the model to integrate task dependencies so that it may represent a context-dependent runtime GM with QoS constraints.

REFERENCES

- [1] E. Yu and J. Mylopoulos, "Why goal-oriented requirements engineering," in *Proceedings of the 4th International Workshop on Requirements Engineering: Foundations of Software Quality*, 1998, pp. 15–22.
- [2] R. Ali, F. Dalpiaz, and P. Giorgini, "A goal-based framework for contextual requirements modeling and analysis," *Requirements Engineering*, vol. 15, no. 4, pp. 439–458, 2010.
- [3] P. Bresciani, A. Perini, P. Giorgini, F. Giunchiglia, and J. Mylopoulos, "Tropos: An agent-oriented software development methodology," *Autonomous Agents and Multi-Agent Systems*, vol. 8, no. 3, pp. 203–236, 2004.
- [4] J. Castro, M. Kolp, and J. Mylopoulos, "Towards requirements-driven information systems engineering: the Tropos project," *Information systems*, vol. 27, no. 6, pp. 365–389, 2002.
- [5] A. Finkelstein and A. Savigni, "A framework for requirements engineering for context-aware services," 2001.
- [6] D. F. Mendonça, R. Ali, and G. N. Rodrigues, "Modelling and analysing contextual failures for dependability requirements," in *Proceedings of the 9th International Symposium on Software Engineering for Adaptive and Self-Managing Systems*, ser. SEAMS 2014. New York, NY, USA: ACM, 2014, pp. 55–64.
- [7] V. R. Basili, G. Caldiera, and H. D. Rombach, "The goal question metric approach," in *Encyclopedia of Software Engineering*. Wiley, 1994.
- [8] V. E. Silva Souza, A. Lapouchnian, W. N. Robinson, and J. Mylopoulos, "Awareness requirements for adaptive systems," in *Proceeding of the 6th International Symposium on Software Engineering for Adaptive and Self-Managing Systems - SEAMS 2011*. New York, New York, USA: ACM Press, May 2011, p. 60.
- [9] V. E. S. Souza and J. Mylopoulos, "From awareness requirements to adaptive systems: A control-theoretic approach," *2011 2nd International Workshop on Requirements@Run.Time*, pp. 9–15, Aug. 2011.
- [10] J. Whittle, P. Sawyer, N. Bencomo, B. H. Cheng, and J.-M. Bruel, "Relax: Incorporating uncertainty into the specification of self-adaptive systems," in *Requirements Engineering Conference, 2009. RE'09. 17th IEEE International*. IEEE, 2009, pp. 79–88.
- [11] L. Baresi and L. Pasquale, "Adaptive Goals for Self-Adaptive Service Compositions," *2010 IEEE International Conference on Web Services*, pp. 353–360, Jul. 2010.
- [12] —, "Live goals for adaptive service compositions," in *Proceedings of the 2010 ICSE Workshop on Software Engineering for Adaptive and Self-Managing Systems - SEAMS '10*, 2010, pp. 114–123.
- [13] F. Dalpiaz, P. Giorgini, and J. Mylopoulos, "Adaptive socio-technical systems: a requirements-based approach," *Requirements Engineering*, vol. 18, no. 1, pp. 1–24, Sep. 2011.
- [14] R. Sebastiani, P. Giorgini, and J. Mylopoulos, "Simple and minimum-cost satisfiability for goal models," in *Advanced Information Systems Engineering*. Springer, 2004, pp. 20–35.